MEME16203Linear Models

Assignment 3

UNIVERSITI TUNKU ABDUL RAHMAN

Faculty:	FES	Unit Code:	MEME16203
Course:	MAC	Unit Title:	Linear Models
Year:	1,2	Lecturer:	Dr Yong Chin Khian
Session:	May 2024		
Due by:	16/7/2024		

Q1. Consider the following models:

$$\text{Model I: } y_{ij} = \beta_0 + \beta_1 x_j + \epsilon_{ij}$$
$$\text{Model II: } y_{ij} = \mu + \alpha_i + \epsilon_{ij}$$

where $i = 1, 2, 3; j = 1, 2; \epsilon_{ij} \sim N(0, \sigma^2)$ for both models and $\mathbf{x_j} = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}^T$

(a) Model I can be written as $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ and Model II can be written as $\mathbf{y} = \mathbf{W}\boldsymbol{\gamma} + \boldsymbol{\epsilon}$. Find matrix $\mathbf{G}$ such that $\mathbf{X} = \mathbf{WG}$. (5 marks)

(b) If $\mathbf{y} \sim N(\mathbf{W}\gamma, \sigma^2\mathbf{I})$, define $\mathbf{P_X} = \mathbf{X}(\mathbf{X^TX})^{-1}\mathbf{X^T}$ for model I and $\mathbf{P_W} = \mathbf{W}(\mathbf{W^TW})^{-}\mathbf{W^T}$ for model II, what is the distribution of $\frac{\mathbf{y}^T(\mathbf{I}-\mathbf{P_X})\mathbf{y}}{\sigma^2}$? Show that the corresponding parameter is a non-centrality parameter? (5 marks)

(c) What is the distribution of $\frac{\mathbf{y^T(I-P_X)P_W(I-P_X)y}}{\sigma^2}$? (10 marks)

(d) You are given that $\frac{\mathbf{y^T(I-P_W)y}}{\sigma^2}$ has central chi-squared distribution with 3 degrees of freedom. Show that $F = \frac{c\mathbf{y^T(I-P_X)P_W(I-P_X)y}}{\mathbf{y^T(I-P_W)y}}$ has an $F-$distribution for some constant c when model II is the correct model. Report a numerical value for c and degrees of freedom. (10 marks)

[Total:30 marks]

Q2. Suppose $\mathbf{Y} = \underset{n\times p}{\mathbf{X}}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I})$ for some unknown $\sigma^2 > 0$. Suppose $\mathbf{X} = \begin{bmatrix} \underset{n\times 2}{\mathbf{X_1}} & \underset{n\times 4}{\mathbf{X_2}} \end{bmatrix}$. Define $\mathbf{P_X} = \mathbf{X}(\mathbf{X^TX})^{-1}\mathbf{X^T}$, $\mathbf{P_1} = \mathbf{X_1}(\mathbf{X_1^TX_1})^{-1}\mathbf{X_1^T}$ and $\mathbf{P_2} = \mathbf{X_2}(\mathbf{X_2^TX_2})^{-1}\mathbf{X_2^T}$.

(a) Determine the distribution of $\frac{1}{\sigma^2}\mathbf{Y^T}(\mathbf{P_X} - \mathbf{P_2})\mathbf{Y}$. (10 marks)

(b) Show that the noncentrality parameter, $\lambda = \frac{1}{\sigma^2}\boldsymbol{\beta}^T\mathbf{X_1^T}(\mathbf{I} - \mathbf{P_2})\mathbf{X_1}\boldsymbol{\beta}$. (10 marks)

(c) Determine the distribution of $\begin{bmatrix} \widehat{\mathbf{Y}} \\ (\mathbf{P_X} - \mathbf{P_1})\mathbf{Y} \end{bmatrix}$. (10 marks)

[Total:30 marks]

Q3. Data were collected to study the effect of temperature on the yield of a chemical process. Two different catalyst $A = -1$ and $B = 1$, were used in the study. Yields were measured under 4 different temperatures for each catalyst. The data are as follows:

Yield (Grams), Y	Temperatures ($^\circ C$), X_1	Catalyst, X_2
y_1	80	-1
y_2	85	-1
y_3	90	-1
y_4	95	-1
y_5	80	1
y_6	85	1
y_7	90	1
y_8	95	1

Consider the model

$$y_i = \beta_1(X_{1i} - 80) + \beta_2 X_{2i} + \epsilon_i$$

where ϵ_i is independently and identically distributed $N(0, \sigma^2)$.

(a) To test for a Catalyst effect, the researchers propose the following test statistic

$$F = \frac{3(\sum_{i=5}^{8} Y_i - \sum_{i=1}^{4} Y_i)^2}{4SSE}$$

Show that this statistic has an F-distribution. Report its degrees of freedom. (15 marks)

(b) With respect to $\boldsymbol{\beta} = (\beta_1 \quad \beta_2)^T$, describe the null hypothesis that can be tested with the F-test in Part (a). What is the alternative hypothesis? (5 marks)

[Total:20 marks]

Q4. Let $\mathbf{Y} \sim N(\mathbf{W}\boldsymbol{\gamma}, \sigma^2 I)$, where

- $\mathbf{W} = \begin{bmatrix}\mathbf{W_1} \ \mathbf{W_2} \ \mathbf{W_3}\end{bmatrix}$,
- $\mathbf{W_1} = \mathbf{1_{10}}$,
- $\mathbf{W_2} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \otimes \mathbf{1_5}$,
- $\mathbf{W_3} = \mathbf{1_2} \otimes \begin{bmatrix} -10 \\ -5 \\ 0 \\ 5 \\ 10 \end{bmatrix}$, and
- $\boldsymbol{\gamma} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix}$

MEME16203Linear Models

(a) Use Cochran's theorem to find the distributions of

- $\frac{1}{\sigma^2}SSE = \mathbf{e^Te} = \mathbf{Y^T(I - P_W)Y}$, where $\mathbf{P_W} = \mathbf{W(W^TW)^{-1}W^T}$
- $\frac{1}{\sigma^2}R(\gamma_1) = \mathbf{Y^TP_{W_1}Y}$ where $\mathbf{W_1} = \mathbf{1}$ is the first column of $\mathbf{W}$ and $\mathbf{P_{W_1}} = \mathbf{W_1(W_1^TW)^{-1}W_1^T}$.
- $\frac{1}{\sigma^2}R(\gamma_2|\gamma_1) = \mathbf{Y^T(P_{W_2} - P_{W_1})}Y$ where $\mathbf{W_2}$ contains the first two columns of $\mathbf{W}$ and $\mathbf{P_{W_2}} = \mathbf{W_2(W_2^TW_2)^{-1}W_2^T}$.
- $\frac{1}{\sigma^2}R(\gamma_3|\gamma_1\gamma_2) = \mathbf{Y^T(P_W - P_{W_2})Y}$.

(10 marks)

(b) Report a formula for the non-centrality parameter of the non-central F distribution of

$$F = \frac{R(\gamma_3|\gamma_1,\gamma_2)}{SSE/7}$$

Use it to the null and alternative hypotheses associated with this test statistic. You are given that: $\mathbf{W^T(P_W - P_{W_2})W]} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 500 \end{bmatrix}$. (5 marks)

(c) Report a formula for the non-centrality parameter of the non-central F distribution of

$$F = \frac{R(\gamma_2|\gamma_1)}{SSE/7}$$

Use it to identify the null and alternative hypotheses associated with this test statistic. You are given that: $\mathbf{W^T(P_{W_2} - P_{W_1})W]} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

(5 marks)

[Total:20 marks]